Wed Nov 07 11:28:19 2001 454

REDACTED

From:
To: "Yamron, Jon" <jyamron@voicesignal.com>
Subject: RE: brainstorm
Date: Tue, 18 Sep 2001 16:27:46 -0400

-----Original Message-----
From: Jon Yamron [mailto:jyamron@voicesignal.com]
Sent: Tuesday, September 18, 2001 3:17 PM
To:
Subject: Re: brainstorm

I sort of see how this might work in the case of white (or similar)
noise.
But what about door slams, lip smacks, telephones, etc.? Or are you not
trying to get at that here?

- Jon

On Tue, 18 Sep 2001, you wrote:
>
> 22nd century data collection
> or
> how to spend gobs of time and money...
> <<permissible noise.doc>>
>
> _____
>
> Voice Signal Technologies, Inc.
> 300 Unicorn Park Drive
> Woburn, MA 01801-3363
>
> Fax: (781) 970-5300
> www.voicesignal.com
>
>
>

Content-Type: application/msword; name="permissible noise.doc"
Content-Transfer-Encoding: base64
Content-Description:

Talk for VST Research, 26 April 2001

References:

[1] S. Martin, J. Liermann, and H. Ney, "Algorithms for bigram & trigram word clustering," Eurospeech '95, Madrid, September 1995.

[2] P.F. Brown, V.J. Della Pietra, P.V. deSouza, J.C.Lai, and R.L. Mercer, "Class-based n-gram models of natural language," Computational Linguistics 18:4, pp. 467-479 (1992).

[3] L. Moisa and E. Giachin, "Automatic clustering of words for probabilistic language models," Eurospeech '95, Madrid, September 1995.

0) Overview of work on class models
   - o What is a class model?
   - o How can we generate classes?
   - o Details of the clustering algorithm
   - o Building class models
   - o Word and class model perplexities
   - o Other clustering algorithms
   - o Other class models

1) What is a class model?
   - o What is a word model?
     Basic discount/backoff models
     Good: Very specific information about common word sequences
     Bad: Big, cutoff required (loss of data), smoothing required (using weaker statistics) to infer data for unseen/discarded events, fragile in case of training/test mismatch
   - o Word classes
     Leverage statistics of common words to induce model of rarer words
     Reduce the number of model parameters
   - o Class model
     Words assigned to a single class
     Good: Can be small, minimal smoothing required, decent statistics used to infer data for unseen events, robust to training/test mismatch
     Bad: Loss of specificity for common word sequences

2)

3)

Here is our first cut at a set of coding tasks that we thought we might accomplish in the next three months, together with possible work assignments:

REDACTED

Manfred/ REDACTED	release process
Jon/ REDACTED	load LMs from file
Jon/	class based LM
Bob/ REDACTED	fix adaptation (use internal silences and alternate prons)

REDACTED

Bob	numerous search issues (thresholding, memory optimization, right generic models, variable (1, 2, 3..) node models, whole word models, etc.
	REDACTED
//Bob	efficient likelihoods (possibly via tables, Laplacian vs Ga
Bob// REDACTED	means and devs
Manfred	grammar based decoding (for Continus Speech Command+Control not neccessarily grammars)
Manfred	dynamic vocabulary (adding, removing words; Master word database)
	REDACTED
All	Documentation
All	error reporting
All	warning removal

REDACTED

Other tasks:

Bug tracking
LM Adaptation (Cache LM certainly)
Tokenization
Perplexity, OOV rates
VSTUtil support
Significance testing
More tests for recognizer
Run Windows Purify
Use HTK feature format. Is it any good? Stackable?
Move to
Run experiments on target hardwar (Emulator)
Pron Guesser.

REDACTED

Subject: ELVIS
Date: Thu, 19 Apr 2001 14:37:48 -0400
From: "Gillick, Larry" <lgillick@voicesignal.com>
To: _Research <research@voicesignal.com>
CC:

REDACTED

The ELVIS project lies at the heart of VST's overall plans for the future. It is clear, however, that ELVIS is a very large undertaking. Not only does it involve the development of highly complex adaptive software, but it also will rely on the creation or acquisition of suitable language corpora (lexicons, text data, acoustic data) and the creation of language and acoustic models. In order for us to successfully carry out this complex project, we must introduce further structure into our organization.

We are very happy to announce that REDACTED has agreed to be the overall ELVIS project coordinator. REDACTED will focus on clearly defining the API for ELVIS as well as the many tasks to be carried out in order to support that API. He will also be coordinating our efforts to complete all of these tasks in a rational and orderly manner.

In order to better support our ELVIS development efforts, we have also identified functional leadership in a number of important technical areas.

Recognizer: REDACTED and Manfred Grabherr
Acoustic modeling: REDACTED
Language modeling and data: Jon Yamron
Front End: REDACTED
Testing and Tuning: REDACTED

We may well augment this list as our plans mature.

Larry